

Influence of roasting level on near infrared spectroscopy prediction of Arabica content from ground coffee blends

Davrieux, F. ^a; Bastianelli, D. ^b; Flori, A. ^a; Manez, J.C. ^a and Guyot, B. ^a

^a CIRAD, TA 80/16, 34398 Montpellier Cedex 5, France. E-mail: davrieux@cirad.fr

^b CIRAD, Laboratoire d'Alimentation Animale, TA 30/A, 34398 Montpellier Cedex 5, France

Keywords: near infrared spectroscopy, coffee blends, roasting, Arabica content

Introduction

Industrialists and distributors are increasingly blending Arabica and Robusta coffees to develop unique coffees for niche markets. This trend corresponds to a real demand for certified products with defined specifications associated with sourcing organically grown beans purchased under fair trading conditions. Consequently, determining the relative proportion of Arabica and Robusta beans in a roasted coffee blend is commercially important.

Chemical analysis based on specific compounds is long, costly and not very reliable for determining the Arabica percentage. Different studies [1, 2] have shown the potential of using near infrared (NIR) spectroscopy to distinguish between Arabica and Robusta coffees. In extending this procedure we have developed a diffuse reflectance NIR spectroscopy model [3] capable of predicting the Arabica content of a blend with a standard error of prediction (SEP) of 2.40%. The accuracy of the predictive model was highly satisfactory provided its basis was robust. This required knowledge of, and the ability to control, the sources of variation. Whilst variability due to the geographical origin of the coffees, the genotype of the beans and post-harvest processing were taken into account in the spectral database, variability due to roasting, the quality of the green coffees due to defects such as the presence of black beans, the presence of foreign bodies such as chickpeas, skins and barley, and the effect of blending to create special niche market coffees remain to be investigated.

In this study, we carried out a specific experiment involving five levels of blending and three levels of roasting (light, medium and dark) to test the influence of roasting level on the SEP estimation of Arabica content, water content and trichromatic coordinates of ground coffee.

Experimental procedure

Reference calibration database

A spectral database of 470 ground samples consisting of pure Arabica and Robusta coffees and proportions of each type blended in the laboratory was assembled following scanning in diffuse reflectance using a FOSS NIRSystem 6500 spectrometer. A mathematical model using multivariate regression methods (partial least squares - PLS) was developed [3] to predict the Arabica content of blends from the NIR spectral data using the spectral segment between 908 and 2,500 nm. The model was validated by predicting a set of 50 samples that were fully independent from the established base. The SEP was estimated to be 2.40% on a weight per unit weight basis with a coefficient of determination (R^2) of the regression of 0.99 and a regression slope of 1.01. This calibration was robust for variability arising from varieties, geographical origins and post-harvest processing. This equation was used to predict the Arabica content in the current experiment.

Experimental design for assessing the roasting factor effect

A two-way factorial design trial was used to assess the effect of three degrees of roasting (light, medium and dark) of five blends involving 0% Arabica and 100% Robusta, 25% Arabica and 75% Robusta, 50% Arabica and 50% Robusta, 75% Arabica and 25% Robusta, 100% Arabica and 0% Robusta coffees. The design was undertaken as two complete replicates.

The Arabica coffee was sourced from Costa Rica and the Robusta coffee was sourced from the Ivory Coast. Each coffee type was roasted separately in a Probat Werke laboratory roaster at 200°C.

The degree of roasting was defined by time as 8 minutes for a light roast, 8 minutes 45 seconds for a medium roast and 9 minutes 40 seconds for a dark roast.

The homogeneity of the different batches was checked by estimating the trichromatic CIE 1976 L^* , a^* , b^* coordinates of beans using NIR spectroscopy calibration equations previously developed in our laboratory and by measuring their loss in weight.

Methods

Sample preparation

After roasting 30 g samples of each blend were prepared by weighing and ground in the laboratory.

Near infrared spectroscopy

Three gram samples of each blend were homogenized and scanned in diffuse reflectance in the spectral range from 400 to 2,500 nm, in 2 nm steps, using a NIRSystem 6500 spectrometer (Foss). The 30 spectral analyses encompassing both replicates were carried out in random order in one hour at constant atmospheric temperature. Spectral data were analysed using WINISI 1.5 (InfraSoft International) software.

Statistical analysis

The effect of blending and roasting on NIR spectroscopic prediction of Arabica content, water content trichromatic coordinates of samples of ground coffee were assessed by analysis of variance. The calculations were carried out with XLSTAT 7.1 software (Addinsoft). The differences between means were tested by the Newman and Keuls test at 5%.

Results

Effect of blending and roasting on Arabica content

Roasting had a significant effect ($P < 0.0001$) on the predicted Arabica content with no significant interaction between blending level and degree of roasting (Figure 1). The average contents were 52.3%, 49.9% and 44.8% for light, medium and dark roasting respectively whereas the true average value for the blends as a whole was 50%. At the 5% level there was no significant difference between replicates.

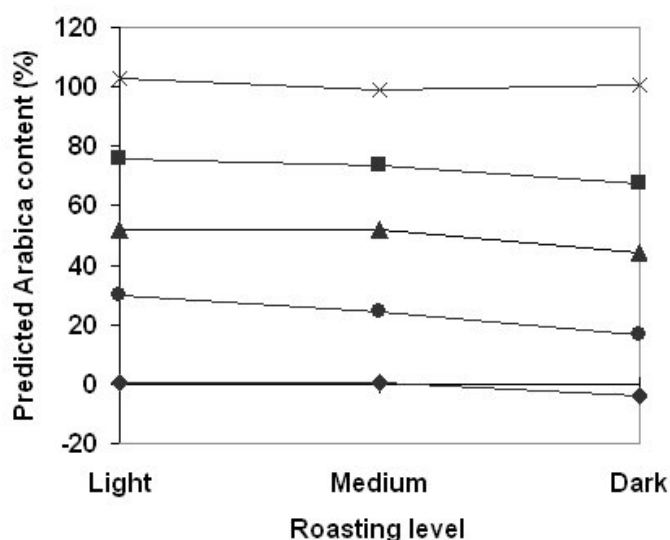


Figure 1. Relationship between mean estimated Arabica content (%) and roasting level (Light, Medium and Dark). Reference levels of Arabica content were X = 100%, ■ = 75%, ▲ = 50%, ● = 25%, ◆ = 0%.

The SEP estimated by comparing all the 30 predicted values 30 with the reference values was 4.21%, with a bias of 1.01, a slope 0.978 and a coefficient of determination of prediction (R^2) of 0.99.

Table 1 shows the statistical criteria for the prediction of Arabica content using the previously derived calibration equations for each roasting level. The SEP for the Dark roasting (6.73%) was significantly higher than the values for the Medium roasting (1.95%) and the Light roasting (2.97%) which were not significantly different.

Table 1. Statistical criteria of comparison between reference and predicted values when estimating Arabica content (%).

Roasting level	SEP	Bias	SEPC	Slope	R^2	Average H
Light	2.97	-2.26	2.04	0.99	0.99	1.54
Medium	1.95	0.13	2.05	1.01	0.99	1.64
Dark	6.37	5.16	3.94	0.95	0.99	3.45

SEP: standard error of prediction

SEPC: standard error of prediction bias corrected

R^2 : coefficient of determination

H: Mahalanobis spectral distance

The Mahalanobis distance (H) values, calculated for the samples by comparison with the reference spectral database, also depended on the degree of roasting (Table 1). These values were all under 3 for Light and Medium roasting, and over 3 for Dark roasting.

The values predicted for light and medium roasting were not significantly different.

Effect of blending and roasting on water content

The water content of the samples was also affected by the degree of roasting, with average levels of 1.20%, 1.05% and 0.97% for Light, Medium and Dark roasting respectively. The blending and block factors did not have any effect on the water content with no significant first-order interaction.

Effect of blending and roasting on trichromatic coordinates

The values of the L, a, b coordinates predicted for Light and Dark roasting corresponded to extreme values when compared to the calibration database, especially in the case of Dark roasting, for which no sample in the database had such low values for each of the three coordinates.

The analysis of variance showed that the roasting factor had a highly significant effect ($P < 0.0001$) on the values of the trichromatic coordinates. The mean luminance (L) values were 23.9, 22.2 and 20.8 for Light, Medium and Dark roasting respectively. The block effects and block x roasting interaction effects were not significant. There was a significant blend effect ($P < 0.012$) with a significant difference between 100% and 0% Arabica content and between 100% and 25% Arabica content. This could be explained by difficulty experienced in carrying out identical roasting between the Arabica and Robusta coffees. The differences were subsequently cancelled out by the blending level.

It could be concluded from these results that roasting had been properly carried out, that the groups were homogeneous and that there were significant differences between the groups. The three roasting levels were well aligned on a similar brown shade, only the luminance or intensity of the brown shade varied from light brown to dark brown.

Discussion

The presented results show that the degree of roasting significantly affected estimates of Arabica content predicted by NIR spectroscopy. This source of variation can be estimated by measuring, or by predicting the trichromatic coordinates of the samples, and by controlling the Mahalanobis distance (H) values at the same time. The trichromatic coordinates in this study were predicted with a model where the ratio of performance to deviation (RPD) exceeded 7. By proceeding in this way, it is possible to identify samples likely to lie outside the norms fixed by the spectral database due to

extremely light or extremely dark roasting. Such samples will have extreme values for some, or all, of the trichromatic coordinates, and H values over 3. In such cases the predictions based on these samples can only be considered as purely indicative.

The significant influence of roasting, even when visible wavelengths apparent to the human eye as different colours are removed from the model, can be explained by the radical biochemical changes that occur inside the product depending on temperature. The presence of water, sugars and amino acids lead to the formation of compounds through thermal degradation [4]. A rise in temperature pushed to the extreme leads to the formation of pyrolysis products. Moreover, the fat in coffee migrates to the surface of the beans in line with temperature [5]. This undoubtedly induces a change in the spectral perception of the fats.

The correction to be made to the model is clearly to add this source of variation to the spectral database. The easiest way, once acceptable roasting limits have been defined in relation with the sensory analysis, is to prepare light and dark roasted pure Arabica and pure Robusta blends and add those samples to the spectral database in order to develop a new prediction model.

Conclusion

In its current form, the predictive model is extremely efficient and covers a wide share of the variability found in commercial coffees, particularly variability due to roasting. The current model particularly demonstrated its potential for recognizing lightly roasted Robusta coffees, with a view to attenuating their typical characters of strength and bitterness, whether in a pure or blended form.

Sources of variation due to the presence of foreign material along with the influence of new cultural practices leading to specific and particular products remain to be investigated.

References

- 1 G. Downey and J. Boussion. *J. Sci. Food Agric.* **71**, 41 (1995).
- 2 R. Briandet, E.K. Kemsley and R.H. Wilson. *J. Agric. Food Chem.* **44**, 170 (1996).
- 3 F. Davrieux, S. Laberthe, J.C. Manez, B. Guyot. *Proceedings of the 19th Conference of the International Coffee Science Association*, Trieste, Italy (2001). <http://www.asic-cafe.org/htm/eng/sectioneng.php?code=ch&number=19>
- 4 J.E. Hodge. *J. Agric. Food Chem.* **1**, 928 (1953).
- 5 A. Illy and R. Viani, in *Espresso coffee: the chemistry of quality*. Academic Press, San Diego, USA, p. 87 (1995).